UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/657,421	09/08/2003	Jordan Cohen	112855.122US2	9023

23483 7590 09/30/2009

WILMERHALE/BOSTON
60 STATE STREET
BOSTON, MA 02109

EXAMINER	
SHAH, PARAS D	

ART UNIT	PAPER NUMBER
2626	

NOTIFICATION DATE	DELIVERY MODE
09/30/2009	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

michael.mathewson@wilmerhale.com
teresa.carvalho@wilmerhale.com
sharon.matthews@wilmerhale.com

Office Action Summary	Application No.	Applicant(s)	
	10/657,421	COHEN ET AL.	
	Examiner	Art Unit	
	PARAS SHAH	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *19 June 2009*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-4, 6-9,12, and 15* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-4,6-9,12 and 15* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

1. This communication is in response to the Arguments filed on 06/19/2009. Claims 1-4, 6-9, 11, 12, and 15 remain pending. The Applicants' remarks have been carefully considered, but they do not place the claims in condition for allowance. Accordingly, this action has been made FINAL.

2. All previous objections and rejections directed to the Applicant's disclosure and claims not discussed in this Office Action have been withdrawn by the Examiner.

Information Disclosure Statement

3. The information disclosure statement (IDS) submitted on 06/19/2009 is in compliance with the provisions of 37 CFR 1.97. Accordingly, the information disclosure statement is being considered by the examiner.

Response to Arguments

4. Applicant's arguments (pages 2-4) filed on 06/19/2009 with regard to claims 1-4, 6-9, 1, 12, and 15 have been considered but are not persuasive for the reasons mentioned below.

With respect to claim 1, the Applicants argue that the secondary reference of Meredith does not align a synthesized word with a spoken utterance as recited since Meredith aligns a phonetic transcription with the pitch measurements in the spoken utterance. The Examiner respectfully disagrees with this assertion. Figure 4B, in step 420 specifically describes the alignment of the spoken utterance (natural) with that of

the synthesized utterance. This alignment is done with respect to the voiced and unvoiced segments. Further, in response to the phonetic transcription being aligned, the phonetic transcription as well as the pitch measurements make up the synthetic speech (see Figure 4A, step 415). Therefore, the alignment is performed in Meredith between the synthesized speech and the spoken utterance, where the synthesized speech is the phonetic transcription and associated voicing characteristics. The claims do not require that the alignment be between the actual speech (audio), but rather the alignment of the word with the spoken utterance. The alignment of word in Meredith is described in col. 4, lines 14-25, where a sample utterance from the user, which is a word, and the same input is used to allow the synthetic voice to have the pitch contour as the natural utterance (see col. 4, lines 26-30). Therefore, the Applicant's argument is not persuasive.

As to the Applicant's second argument asserting that Meredith receives two inputs and Marasek only receives one input and therefore Meredith's technique cannot be applied to Marasek's system. The Examiner respectfully disagrees with this assertion. The test for obviousness is not whether the features of a secondary reference may be bodily incorporated into the structure of the primary reference; nor is it that the claimed invention must be expressly suggested in any one or all of the references. Rather, the test is what the combined teachings of the references would have suggested to those of ordinary skill in the art. See *In re Keller*, 642 F.2d 413, 208 USPQ 871 (CCPA 1981). Further, the combined teachings of Marasek in view of Meredith would have been obvious to one of ordinary skilled in the art as the

modification to Marasek's system with that of Meredith enable the synthesized speech utterance to have a more natural intonation based on a user inputted utterance (see Meredith col. 2, lines 58-64) thus having a more pleasing speech output. Therefore, the Applicant's argument is not persuasive.

With respect to claim 1, the Applicant's argue that none of the prior art teaches or suggests the sequence of event that are claimed. Applicant's arguments fail to comply with 37 CFR 1.111(b) because they amount to a general allegation that the claims define a patentable invention without specifically pointing out how the language of the claims patentably distinguishes them from the references. The Applicant alleges that that cited references do not teach the following sequence but does not specifically point out how the sequence is not taught by the combination of references. Therefore, the Applicant's argument is not persuasive.

With respect to claim 1, Applicant's argue that one skilled in the art would not have combined Marasek's system with that of Cameron for implementation on a handheld device. The Examiner respectfully disagrees with this assertion. Although it is true that Marasek's system is used for generating synthesized speech with personality patterns, Marasek in the disclosure does not teach away or against the system being implemented on a handheld device. In Marasek, page 2, paragraphs [0007]-[0008] and paragraphs [0016], Marasek teaches the implementation in a man-machine interface dialogue system. Similarly, in Cameron, page 18, sect. 6, specifically describes such an interface where the portable device dials a number where the user first inputs speech and the voice assistant prompts the user in a man-machine dialogue manner in order to

obtain the correct entry. Therefore, the implementation of Marasek's system into a handheld device would have been obvious to one of ordinary skilled in the art as taught in Cameron for the portability nature of handheld devices and to further promote a handsfree/eyes-free environment (see Cameron, page 3). Therefore, the argument is not persuasive.

In response to applicant's argument that the examiner's conclusion of obviousness is based upon improper hindsight reasoning, it must be recognized that any judgment on obviousness is in a sense necessarily a reconstruction based upon hindsight reasoning. But so long as it takes into account only knowledge which was within the level of ordinary skill at the time the claimed invention was made, and does not include knowledge gleaned only from the applicant's disclosure, such a reconstruction is proper. See *In re McLaughlin*, 443 F.2d 1392, 170 USPQ 209 (CCPA 1971). It is further pointed out by the Examiner that the each combined reference is supplied with a proper motivation which further supports why one skilled in the art would have combined each of the prior art references to obtain the invention as claimed and have been properly provided in the prior issued Office Action. Therefore, the argument is not persuasive.

Claim 9 and claimed dependent upon claim 1 are rejected for similar reasons as mentioned above.

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

6. Claims 1-3, 6-7, 9, 12, and 15 are rejected under 35 U.S.C. 103(a) as being unpatentable over Marasek *et al.* (EP 1 271 469) in view of Meredith (US 5,796,916, issued on 08/18/1998) in view of Lumelsky (US 6,081,780, issued on 06/27/2000) in view of Cameron (WO 02/097590 A, published on 12/05/2002).

As to claims 1and 9, Marasek teaches a method and system for speech synthesis comprising:

receiving a spoken utterance (see Abstract Figure 1, S1, receive speech input S1) (e.g. It is obvious that a microphone is used to input speech in the system.);

in response to receiving a spoken utterance (see Figure 1, S1, speech is received and corresponding processing performed as shown in the steps below S1):

extracting one or more prosodic parameters (see Figure 1, S21m, extract prosodic features) (e.g. It is obvious that a signal processor would be used to extract prosody features such as described un [0042] and is well known in the art.) from the spoken utterance;

performing speech recognition on the spoken utterance to generate a recognized word (see Figure 1, steps S12 and [0040] recognize speech)

generating a prosodic mimic word using (Figure 1, step S40 and S50 and [0046], speech synthesis is performed on the input speech by applying prosody to a given text (see [0003]) and the one or more prosodic parameters (see Figure 1, step S21, prosody parameters are extracted and applied to storage personality pattern as seen in Figure 1, step S30).

However, Marasek does not specifically teach the alignment of the spoken utterance and the synthesized word.

Meredith does disclose the alignment of the spoken utterance to the synthesized speech (see Abstract).

It would have been obvious to one of ordinary skilled in the to art at the time the invention was made to have combined the speech synthesis for an utterance as presented by Marasek by the alignment of the utterance and the synthesized word presented by Meredith. The motivation to have combined the two references includes the improvement in intonation (see Meredith col. 3, lines 5-10).

However, Marasek in view of Meredith do not specifically teach the generation of a nominal word.

Lumelsky does teach synthesizing a nominal word (e.g. The applicant refers to the nominal word as synonymous to synthesized word) corresponding to the recognized word (see col. 13, lines 29-41, synthetic speech is produced

based on a pre-stored voice selected by the narrator. Further in col. 16, lines 45-65, the speech that has been output can be reconfigured by editing or changing the prosody parameters.); and

It would have been obvious to one of ordinary skilled in the to art at the time the invention was made to have combined the speech synthesis for an utterance as presented by Marasek in view of Meredith by the generation of a default voice output as taught by Lumelsky. The motivation to have combined the references involves editing or altering of output based on user preference (see Lumelsky, col. 16, lines 42-65).

However, Marasek in view of Meredith in view of Lumelsky do not specifically disclose the system implemented on a handheld device and at least one of a command to be executed by the handheld device and a name to be dialed by a handheld device and if the recognized word includes the command, executing the command on the handheld device, and if the recognized word includes the name, dialing a number corresponding to the name.

Cameron does disclose the speech synthesis implemented on a handheld device (see page 5, 6th paragraph and see page 29, 1st paragraph) (e.g. portable is synonymous to handheld and PDA is a handheld device)

at least one of a command to be executed by the handheld device (see page 18, sect. 6, first paragraph, lines 1, command dial makes voice assistant to dial the telephone) and a name to be dialed by a handheld device (see page 18, sect. 6, 2nd paragraph, lines 5) and;

if the recognized word includes the command, executing the command on the handheld device (see page 18, sect. 6, first paragraph, lines 1, command dial makes voice assistant to dial the telephone), and if the recognized word includes the name, dialing a number corresponding to the name (see page 18, sect. 6, 2nd paragraph, lines 5-9 calls John at business number using autodial).

It would have been obvious to one of ordinary skilled in the to art at the time the invention was made to have combined the speech synthesis for an utterance as presented by Marasek in view of Meredith in view of Lumelsky by the implementation on a handheld device for the purpose of portability, which allows the user to use the device anywhere as is apparent and seen in navigation and translation devices, which incorporate speech recognition and generate a synthetic speech output based on user selection (see Cameron page 5, last paragraph, example of recognition and voice output is described and page 10, bullet 10-page 11, command recognition and speech synthesis) for minimal hand/eye distraction (see Cameron, page 32, 2nd paragraph).

As to claim 9, Marasek in view of Meredith in view of Lumelsky in view of Cameron teaches all of the limitations as in claim 1, above. Furthermore, Cameron teaches the use of a processor in for the recognition of command and dialing a number (see Figure 1, CPU)

As to claim 2, Marasek in view of Meredith in view of Lumelsky in view of Cameron teaches all of the limitations as in claim 1, above.

Furthermore, Marasek teaches wherein the one or more prosodic parameters include pitch (see [0042], pitch).

As to claim 3, Marasek in view of Meredith in view of Lumelsky in view of Cameron teaches all of the limitations as in claim 1, above

Furthermore, Marasek teaches wherein the one or more prosodic parameters include timing (see [0042], speech element duration).

As to claim 4, Marasek in view of Meredith in view of Lumelsky in view of Cameron teaches all of the limitations as in claim 1, above.

Furthermore, Marasek teaches wherein the one or more prosodic parameters include energy (see [0042], loudness).

As to claim 6, Marasek in view of Meredith in view of Lumelsky in view of Cameron teaches all of the limitations as in claim 1, above

Furthermore, Meredith teaches comprising temporally (see col. 4, lines col. 4, lines 37-53) (e.g. The reference indicates the use of intervals and a pitch point marking) aligning phones (see col. 3, line 5) (e.g. Phones are synonymous to phonetic symbols) of the spoken utterance and phones of the nominal word (see Abstract).

As to claim 7, Marasek in view of Meredith in view of Lumelsky in view of Cameron teaches all of the limitations as in claim 1, above

Furthermore, Marasek teaches comprising converting the prosodic mimic word into a corresponding audio signal (see Figure 1, steps S40 and S50 and [0048], synthetic speech is output) (e.g. It is obvious that the signal is in audio form in order for the user to listen to the speech generated).

As to claim 12, Marasek in view of Meredith in view of Lumelsky in view of Cameron teaches all of the limitations as in claim 9, above

Furthermore, Lumelsky teaches a storage device (see col. 17, line 22, dsp) including executable instructions (see col. 17, line 21) for speech analysis and processing (see col. 17, lines 17-20, dsp).

As to claims 8, Marasek in view of Meredith in view of Lumelsky in view of Cameron teaches all of the limitations as in claim 1, above

Furthermore, Cameron teaches the use of a portable telephone (see page 5, paragraph 6, line 4) input device (see page 5, paragraph 6, line 1) and the prosodic mimic word (synthesis and presentation of commands to the user) (see Abstract and page 18, paragraph 2, lines 1-8) is provided to a telephone output device (see page 5, paragraph 6, line 2). Further, Cameron discloses the use of a user interface (see Abstract) utilizing a mobile phone (see page 18, line 7 and

page 5, paragraph 5, line 4) (e.g. It is inherent that a portable telephone encompasses a mobile telephone).

As to claim 15, Marasek in view of Meredith in view of Lumelsky in view of Cameron teaches all of the limitations as in claim 1, above

Furthermore, Cameron teaches wherein the command is any one of a plurality of available commands (see page 18, sect. 6, first paragraph, lines 1, command dial makes voice assistant to dial the telephone and see page 10, sec. 10, lines 1, where another command search is described and see page 11, bullet c, lines 2, temporal command, hence plurality of commands are utilized)

Conclusion

7. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of

the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

8. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Akamine et al. (US 6,161,091) is cited to disclose speech recognition-synthesis storing prosody information. Freedman (US 7,124,082) is cited to disclose speech-to-text-to speech system. Blass (US 7,280,968) is cited to disclose generation of speech responses including prosodic characteristics. Tang et al. (US 2002/0173962) is cited to disclose speech personalization from text. Kamai (Us 2005/0125227) is cited to disclose speech synthesis from text using spoken utterance. Aaron et al. (US 2005/0071163) is cited to disclose text to speech synthesis using spoken input and alignment.

9. Any inquiry concerning this communication or earlier communications from the examiner should be directed to PARAS SHAH whose telephone number is (571)270-1650. The examiner can normally be reached on MON.-THURS. 7:30a.m.-4:00p.m. EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David Hudspeth can be reached on (571)272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/David R Hudspeth/
Supervisory Patent Examiner, Art Unit 2626

/P. S./
Examiner, Art Unit 2626

09/24/2009